SMS design and aberrations theory

Fabio Corrente[a,b], Pablo Benítez[b], Wang Lin[b], Juan Carlos Miñano[b] and Fernando Muñoz[c]

[a]ERCIM - Alain Bensoussan Fellow

[b]CeDInt - UPM, Campus de Montegancedo 28223 Pozuelo de Alarcón, Madrid - España

[c]General Electric España

ABSTRACT

The SMS, Simultaneous Multiple Surfaces, design was born to Nonimaging Optics applications and is now being applied also to Imaging Optics. In this paper the wave aberration function of a selected SMS design is studied. It has been found the SMS aberrations can be analyzed with a little set of parameters, sometimes two. The connection of this model with the conventional aberration expansion is also presented. To verify these mathematical model two SMS design systems were raytraced and the data were analyzed with a classical statistical methods: the plot of discrepancies and the quadratic average error. Both the tests show very good agreement with the model for our systems.

1. INTRODUCTION

In imaging designs, a generalized Cartesian Oval has long been found to be effective in correcting on-axis spherical aberration.[1,2] An asphere can be constructed directly by fulfilling a constant optical path length from an on-axis image point to an on-axis object point. The SMS (Simultaneous Multiple Surface) method is a direct extension to the Cartesian Oval calculation. Given n input one-parameter ray-bundles and n output one-parameter ray-bundles, n aspheres can be constructed directly. The SMS method was developed primarily as a design method in Non-imaging Optics during the 1990s.[3] The method was then extended for designing Imaging Optics. Classic imaging design methods depend heavily on multi-parametric optimization techniques. The SMS method provides an effective way of overcoming local minimums. The direct construction of n aspheres helps to reduce dramatically the total number of parameters, thus avoiding the appearance of many undesired local minimums. An SMS system is designed imposing the wave aberration functions vanishes for certain set of rays. The aberrations of SMS-3M (i.e. designed using three meridional ray bundles) and SMS-1M2S (i.e. designed by one meridional ray bundle and two skew ray bundles) optical systems with rotational symmetry are described. From a mathematical point of view the SMS design provides a method to write the aberration function in a form similar to what reported in appendix A with two functions multiplying the auxiliary functions g, h.

2. WAVE ABERRATION FUNCTION OF SMS DESIGNS

2.1 SMS-3M design algorithm

The SMS method can be implemented to design using meridian and skew ray-bundles. In the SMS-3M three meridian ray-bundles are used to design three aspheric surfaces. The standard SMS procedure for imaging design is monochromatic, and consists of two steps: selection of the central segments of the surfaces and recursive generalized Cartesian oval calculations.[4] It is illustrated in Fig 1. Such a system is aberration-free for some values, in particular[5] it is $W(\boldsymbol{r}, 0, \theta) = W(\boldsymbol{r}_0, \boldsymbol{\rho}, 0) = W(\boldsymbol{r}_0, \boldsymbol{\rho}, \pi) = 0$ so the aberration function must vanish for these values; $\boldsymbol{r}$ and $\boldsymbol{\rho}$ are the vectors indicating the intersection of the ray on the object and on the pupil plane and θ indicate the angle between them. As indicated in appendix A and other sources[5] this states a functional equation on W whose solution is

$$W(\boldsymbol{r}, \boldsymbol{\rho}, \theta) = (\boldsymbol{r} - \boldsymbol{r}_0)a(\boldsymbol{r}, \boldsymbol{\rho}, \theta) + \sin\theta b(\boldsymbol{r}, \boldsymbol{\rho}, \theta), \tag{1}$$

Corresponding author: fcorrente@cedint.upm.es

Optical Systems Design 2012, edited by Pablo Benítez, et al., Proc. of SPIE Vol. 8550, 855010 · © 2012 SPIE
CCC code: 0277-786/12/$18 · doi: 10.1117/12.981163

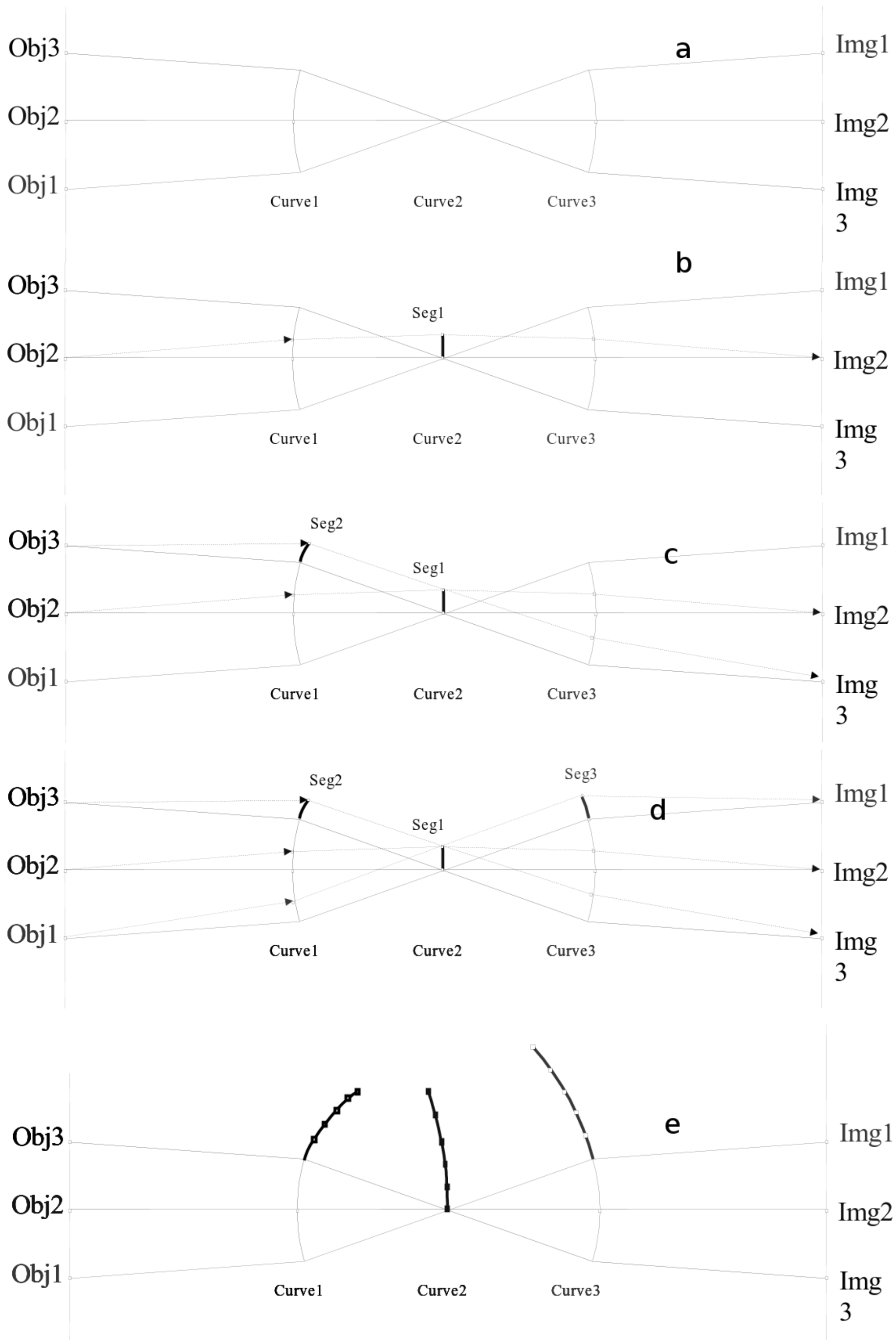

Figure 1. The sms process for a system designed by three meridional ray bundles.

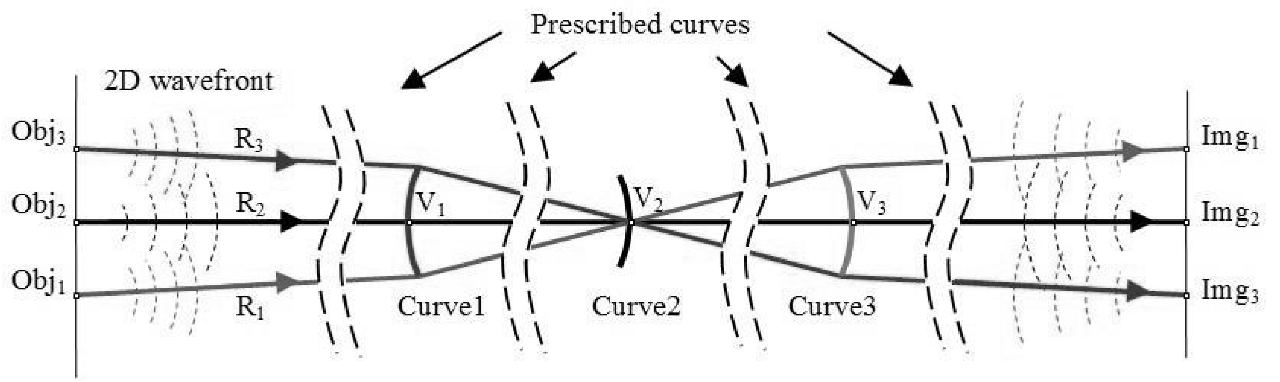

Figure 2. The three meridian ray bundles selected for the SMS design correspond to the rays emitted from 3 object points placed symmetrically about the optical axis, as here shown. V1, V2 and V3 are the central vertices of the initial curves that can be defined beforehand by the designer as a degree of freedom. Consider the rays R1, R2 and R3 shown, which all pass through the vertex of Curve 2: V2. R2 is the on-axis ray emitted from the on-axis point Obj2 to reach its on-axis image point Img2. R1 is emitted from an off-axis point Obj1 to reach its prescribed off-axis image point Img1 and R3 is its symmetrical counterpart.

This function has not explicit rotational symmetry i.e. a and b are not arbitrary. As it is well know[6] this can be obtained using the variables (u, v, w), where $u = \boldsymbol{r}^2, v = \boldsymbol{\rho}^2, w = \boldsymbol{r} \cdot \boldsymbol{\rho}$. According to appendix A then W has the form:

$$W(u, v, w) = w(u - u_0)A(u, v, w) + (uv - w^2)B(u, v, w), \tag{2}$$

It is only needed to expand also A and B and equating the coefficients for every power of $u^m v^n w^t$. Following for example[7],[9] the aberration function is written as power of $\boldsymbol{r}, \boldsymbol{\rho}, \cos\theta$: here it is proposed a different notation writing the expansion to write W as function of u, v, w. Using (u, v, w) the expansion takes the form:

$$W(u, v, w) = \sum_{mnt} C_{mnt} u^m v^n w^t. \tag{3}$$

The null coefficients of the expansions are simply the one with $m = 0$ in this notation. If it is considered the power expansions of A and B the following equations expresses the aberration coefficients:

$$C_{mnt} = A_{(m-1)n(t-1)} - u_0 A_{mn(t-1)} + B_{(m-1)(n-1)t} - B_{mn(t-2)}. \tag{4}$$

Note that the low order aberrations are given by the function A: this is due to $v(u - u_0)$ factor which has lower degree than the factor $(uv - w^2)$ that multiply B. In particular it is used the above model to describe the aberration function of an SMS optical system with different approximation levels.

2.2 One meridional and two skew rays: SMS-1M2S

To obtain the aberration function of this kind of systems it was followed the same way used for the other two. A suitable function could be:

$$W(\boldsymbol{r}, \boldsymbol{\rho}, \theta) \backsimeq \rho[(\boldsymbol{r} - \boldsymbol{r}_0)A(\boldsymbol{r}, \boldsymbol{\rho}, \theta) + \sin\frac{3\theta}{2} B(\boldsymbol{r}, \boldsymbol{\rho}, \theta)], \tag{5}$$

Squaring up every single terms and after some algebra for the factor $\sin(3\theta/2)$ it is obtained:

$$W(u, v, w) \backsimeq v[(u - u_0)A(u, v, w) + \frac{1}{2}(uv^{3/2} + 3uvw - 4w^3)B(u, v, w)]. \tag{6}$$

For the last equation more explanations are needed: first of all note that the term in the second parenthesis $uv^{3/2}/2$ it is not present in the general W expression. In fact this is an expansion in power of u, v, w and can not contain powers with no integer exponent. This strange exponent come out for dimensional analysis: to obtain the same dimension of w^3 and without introducing a $\cos\theta$ factor (remember $w = r\rho\cos\theta$) the term $uv^{3/2}$ must be chosen to multiply a factor 1/2 appearing when it is written $\sin(3\theta/2)$ as function of (u, v, w). One explanation is that the function $B(u, v, w)$ could not be analytic but this is required just for the function W: in this way the product of should be analytic. Moreover take into account that an SMS-1S2M system has a small area in the center of the pupil unused $\rho \leq \rho_{min}$ so it is not aberration free exactly on the lines $\theta = 0, 120, 240$ grades but for the last two values the aberration free lines are parallel to those used in our model. This model can be considered as semi-empirical.

3. TEST PROCEDURE

To test our models the well corrected SMS system described[10] for the 3M design is used. Simulations were run for both the systems to obtain the optical path length differences as function of the object angle and the intersection point on the pupil, then these data were fitted with the function Eq. 2 with different approximations for A and B from constant up to polynomials of degree 5 in u, v, w. The differences between the values predicted by our fitted model and the data of simulation were plotted in Fig. 4.1 and to have an estimation of the mean error of our model the following relation was used:

$$s = \left[\frac{\sum_i^N (W_i -_e W_i)^2}{N - m} \right]^{1/2} \tag{7}$$

where N is the number of point used for the fit, m is the number of parameters computed by the fit, W_i is the ith computed value of W and ${}_eW_i$ is the expected value. The term m at denominator needs some explanation: in a fit some parameters are calculated and more the calculated parameters are more it is easy the function fits them; the limit case is when the computed parameters are equal to the number of data, in this case it is an interpolation, and of course the error it would be zero but this is meaningless. The m term at denominator is needed to correct this anomaly: when the number of parameters grows up the denominator goes down and the error takes into account the big number of computed parameters.[12]

4. RESULTS AND CONCLUSIONS

4.1 Analysis of different orders of approximation

A↓ B→	0	1	2	3	4	5
0	0.09	0.08	0.07	0.06	0.06	0.05
1	0.03	0.006	0.005	0.005	0.004	0.004
2	0.03	0.002	0.002	0.002	0.002	0.002
3	0.02	0.002	0.002	0.002	0.002	0.001
4	0.02	0.002	0.002	0.002	0.001	0.001
5	0.02	0.002	0.002	0.001	0.001	0.001

(b) SMS-1M2S

A↓ B→	0	1	2	3	4	5
0	0.2	0.14	0.13	0.11	0.1	0.1
1	0.04	0.03	0.03	0.02	0.02	0.02
2	0.02	0.01	0.01	0.01	0.008	0.007
3	0.02	0.01	0.01	0.01	0.008	0.007
4	0.01	0.004	0.005	0.004	0.003	0.003
5	0.007	0.003	0.003	0.003	0.002	0.002

Table 1. The discrepancies in order of λ at different order of approximation for an SMS-3M design. In the first row and in the first column are reported the order of expansion of A and B.

Here there are the obtained results of tests for our model. An SMS-3M design was simulated and the values of W for different values of (u, v, w) were obtained. Then the fit of this set of data with the proposed model was run starting from the zero order (A and B constant) up to fifth order in (u, v, w) with the following results. At zero order a good accordance of our model and the data was obtained: the average error Eq. 7 of 0.09λ. Then different polynomial approximations for A and B were used, up to fifth order in (u, v, w), to obtain an average error of 0.001λ. In this case is also presented the same analysis for the SMS-1M2S system, the same

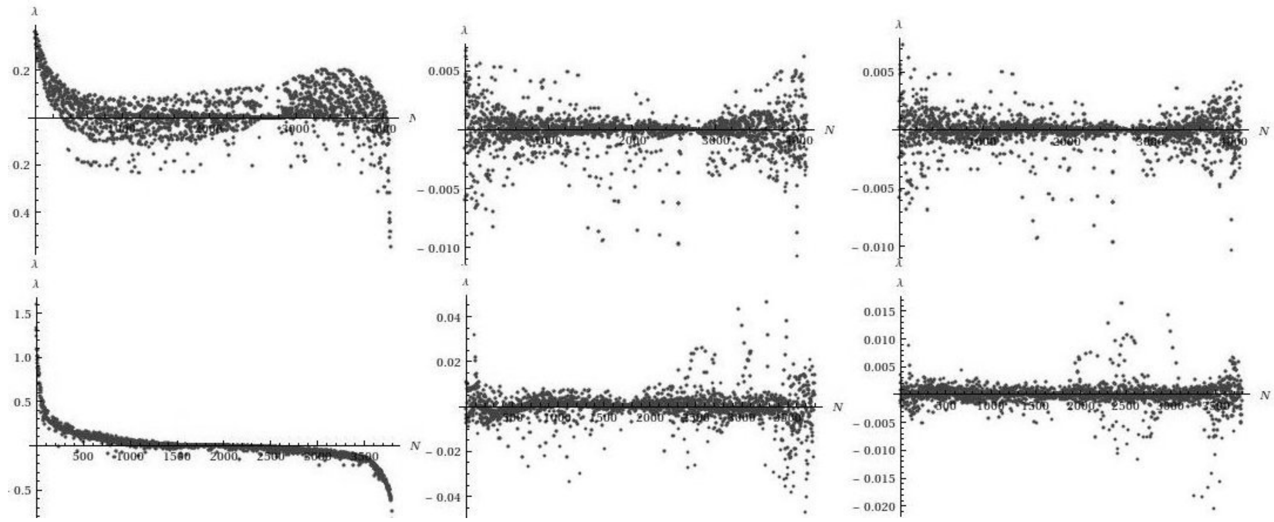

Figure 3. The plotted discrepancies between the proposed models and the data obtained by simulations, on the first line for an SMS-3M design and on the second for an SMS-1S2M. From left to right the order of approximation for A and B are $0, 3, 5$. On the horizontal axes there is just an index indicating the Nth point of the chose for the simulation, on the vertical axes there is the discrepancies. Closer to 0 is better.

considerations subsist also for this system. The same performed analysis for a SMS-3M system is reported in table 4.1 for a SMS-1M2S system. Note that it is always needed to approximate the A function at least with a function linear in the (u, v, w) to have good results.

4.2 Evolution of A and B in an optimization process

Moreover the evolution of the constants A and B during the step of the optimization process was studied. B coefficient varies very much compared with A: in the various steps it decrease up to 95% while A varies of 21%, moreover the B coefficient decreases its value while the A increases. Taking into account for an aberrations free system both A and B vanish this could be explained in this way: main contribution to aberrations is given by the second term and the optimization program try to minimize this one. The variation of the A coefficient is due to the dependence of both the coefficients by the same building parameters. All the results are summarized in the table Eq. 4.2.

Step	A (μm)	B (μm)
1	2259940	64814
2	2465400	39879.5
3	2594210	19969.9
4	2664660	2725.6

Table 2. The evolution of the coefficients A and B during the process of optimization of the lens. There is parabolic surface as first element and its curvature varies from 0 (step 1) up to 0.0006 (step 4) and they are the reported.[10]

4.3 Conclusions and further developments

The mathematical hypothesis done with Eq. 2 are in a very good agreement with the run simulation: this permits to write the aberration function of an SMS system in a simpler and faster way. In fact the identification of the aberration of the non vanishing coefficients can be made a priori and this permits to use the conventional aberration theory to compute them. A very good approximation is obtained with order 2 or 3 approximation for A and B: at this order of approximation still few coefficients are needed to be computed but excellent accuracy is obtained. Going over this orders is for high end design or for applications requiring high level precision. Moreover this mathematical approach can be generalized to other SMS designs or to a generic xMyS design i.e. with x meridional rays and y skews. For an SMS-1M2S design the approximation of A and B with constants is still possible, the mean discrepancy is within 0.2λ i.e. smaller than the 0.25λ universally considered as good tolerance, but better results are obtained if they are used a linear approximation for A and B as constant: in this way the accuracy is improved of an order of magnitude with a small increase of computed quantities (from 2 coefficients to 5).

5. ACKNOWLEDGMENTS

This work was carried out during the tenure of an ERCIM "Alain Bensoussan" Fellowship Programme. This Programme is supported by the Marie Curie Co-funding of Regional, National and International Programmes (COFUND) of the European Commission.

APPENDIX A. APPENDIX: SOLUTION TO FUNCTIONAL EQUATION

If an analytic function $f(x)$ vanishes for a certain value x_0 it can be written as $f(x) = (x-x_0)g(x)$. This descend directly from its series expansion and it can be proofed just expanding the function and collecting its terms. This argument can be easily generalized to functions of several variables.[14] Here it is proofed for an analytic three variables functions that

$$f(x_0, y_0, z_0) = 0 \Leftrightarrow f(x) = (x - x_0)g(x, y, z) + (y - y_0)h(x, y, z) + (z - z_0)l(x, y, z). \tag{8}$$

The proof from right to left is obvious. To proof the reverse the function can be expanded as

$$f(x, y, z) = \sum_{i=1}^{\infty}\sum_{j=1}^{\infty}\sum_{k=1}^{\infty} f_{ijk}(x - x_0)^i(y - y_0)^j + (z - z_0)^k, \tag{9}$$

where the term $f_{000} = 0$ must vanish, in fact $f_{000} = f(x_0, y_0, z_0) = 0$ by hypothesis. If the terms are collected separating the pure powers of x, y, z we have

$$f(x, y, z) = \sum_i f_{i00}(x - x_0)^i + \sum_j f_{0j0}(y - y_0)^j +$$
$$\sum_k f_{00k}(z - z_0) + \sum_i \sum_j \sum_z f_{ijk} f_{ijk}(x - x_0)^i(y - y_0)^j(z - z_0)^k.$$

The last term can be split between the other three, to have:

$$f(x) = \sum_i f_{i00}(x - x_0)^i + \tilde{g} + \sum_j f_{0j0}(y - y_0)^j + \tilde{h} + \sum_k f_{00k}(z - z_0) + \tilde{l} \tag{10}$$

and then rewritten as

$$f(x) = (x - x_0)g(x, y, z) + (y - y_0)h(x, y, z) + (z - z_0)l(x, y, z). \tag{11}$$

The extension of this theorem to more general form

$$f(x, y) = G(x, y)g(x, y) + H(x, y)h(x, y) \tag{12}$$

with G and H analytic is obvious, it is just needed take into account the Taylor expansion of G and H.

REFERENCES

[1] Buchdahl H. A., [An Introduction to Hamiltonian Optics] Cambridge University Press, (1970)

[2] Born M. Wolf E., [Principles of Optics], Cambridge University Press (1997)

[3] Winston R. , MiñanoJ. C. and Benítez P., with contributions of Shatz N., BortzJ., [Nonimaging Optics], Academic Press Elsevier, Chap. 8 (2004).

[4] Miñano J. C. , Benítez P., Wang L., Infante J., Muñoz F., and Santamaría A., "An application of the SMS method for imaging designs," Opt. Express 17(26), 24036–24044 (2009).

[5] Muñoz Fernández F., Sistemas ópticos avanzados de gran compatibilidad con applicaciónes a formación de imagen y iluminación, PhD thesis, Madrid (2004)

[6] Malacara D., Malacara Z., [Handbook of optical design], Marcel Dekker, Inc. (2004)

[7] Schulz G., "Higher Order aplanatism", Optics communications, (1982), 41(5) p 315 -319

[8] Schulz G., "Aberration-free imaging with large fields with thin pencils", Optica Acta (1985), 32(11) p 1361 - 1371

[9] Schulz G., "Imaging performance of aspherics in comparison with spherical surfaces", Appl. Opt. 26, 5118-5124 (1987)

[10] Wang L. et alt., "SMS-based optimization strategy for ultra-compact SWIR telephoto lens design" Optics Express, Vol. 20, Issue 9, pp. 9726-9735 (2012)

[11] Mahajan V.N., [Aberration Theory made simple] Bellingham, Washington SPIE Optical Engineering Press, (1991)

[12] Taylor J.R. [An Introduction to Error Analysis] Sausalito, California, University Science Books, (1997)

[13] Wang L., Benítez P., Miñano J. C., Infante J. and Biot G., "Progress in the SMS design method for imaging optics", Proc. SPIE 8128, 81280F (2011)

[14] Castillo E., Iglesias A., Ruíz-Cobo M.R., Functional Equations in Applied Sciences, (Elsevier, 2005)